| OF |
AD
A067187

END
DATE
FILMED
6-79
DDC
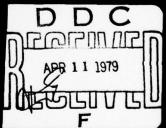
LEVEL II

The Florida State University
Department
of
Statistics
Tallahassee, Florida
32306

6

MINIMAX SUBSET SELECTION FOR THE MULTINOMIAL AND POISSON DISTRIBUTIONS.

by

Roger L. Berger
The Florida State University

FSU STATISTICS REPORT M497
9 Technical Report USARO-D33

12 25 p.

14 FSU-STATISTICS-M497, ARO-D33

The Florida State University
Department of Statistics
Tallahassee, Florida 32306
11 February, 1979

AMS 1970 subject classifications. Primary 62F07; Secondary 62C05.

Key words and phrases. Linear programming, minimax subset selection, expected subset size.

400 277

ABSTRACT

Let $(X_1, \ldots, X_k)$ be a multinomial vector with unknown cell probabilities $(p_1, \ldots, p_k)$. A subset of the cells is to be selected in a way so that the cell associated with the smallest cell probability is included in the selected subset with a preassigned probability, P*. Suppose the loss is measured by the size of the selected subset, S. Using linear programming techniques, selection rules can be constructed which are minimax with respect to S in the class of rules which satisfy the P*-condition. In some situations, the rule constructed by this method is the rule proposed by Nagel (1970). Similar techniques also work for selection in terms of the largest cell probability. The rules constructed in this fashion are also minimax for selection in terms of Poisson parameters.

1.   INTRODUCTION

In this paper, subset selection problems for the multi-nomial distribution are considered. In these problems the aim is to select a non-empty subset of the cells which contains the cell with the highest or lowest cell probability. The goal is to find a selection rule which includes the highest or lowest cell probability with probability at least equal to a preassigned number, P^*. Having satisfied this minimum requirement, the goal is to find a rule which effectively excludes non-best cells. This leads to the use of the number of cells selected or the number of non-best cells selected as a measure of the loss to the experimenter. Minimax rules for these losses are considered and the main result of this paper is that minimax rules can be constructed by solving the appropriate linear programming problem. The rule obtained in this way in some situations corresponds to a particularly simple and easy to implement rule proposed and studied by Nagel (1970). This rule can also be shown to have another optimal property if the cell probabilities are in a slippage configuration.

The subset selection problem for multinomial distributions has been previously considered by Gupta and Nagel (1967), Nagel (1970), Panchapakesan (1971) and Gupta and Huang (1975). Minimax subset selection rules have been recently investigated in a general setting by Berger (1979).

Section 2 contains the necessary notation for a formulation of the problem. In Section 3, the problem of choosing the smallest cell probability is considered and the fact that mini-max selection rules can be constructed using linear programming methods is proven. Examples of rules constructed in this way are given. A particular rule which arises as the solution to the problem is further considered in Section 4. This rule is found to have a certain optimality property if the parameters are in a slippage configuration. In Section 5, the analogous results for the problem of selecting the largest cell proba-bility are outlined. In Section 6, the fact that the rules constructed in Sections 3 and 5 are also minimax for selection in terms of Poisson parameters is explained.

2. NOTATION AND FORMULATION

Let $\underset{\sim}{X} = (X_1, \ldots, X_k)$ be a multinomial random vector with $\sum_{i=1}^{k} X_i = N$. Let $\underset{\sim}{p} = (p_1, \ldots, p_k)$ be the unknown cell probabilities with $\sum_{i=1}^{k} p_i = 1$. The ordered cell probabilities will be denoted by $p_{[1]} \leq \ldots \leq p_{[k]}$. The goal of the experi-menter is to select a subset of the cells including the __best__ cell, the cell associated with $p_{[1]}$. A __correct selection__, CS, is the selection of any subset which contains the best cell. If for a particular parameter value, $\underset{\sim}{p}$, more than one of the cells is tied with the smallest p_i, one of these cells will be considered to be tagged and a CS occurs if the tagged cell is selected. This assumption is not of essential importance and can be dropped without effecting any of the results. But it

eliminates arguments involving limits of parameter values.

A selection rule will be denoted by $\phi(x) = (\phi_1(x), \ldots, \phi_k(x))$ where $\phi_i(x)$ is the probability of including the ith cell in the selected subset having observed $X = x$. The $\phi_i(x)$ are called the _individual selection probabilities_. To insure that a non-empty subset is always selected, only rules which satisfy $\sum_{i=1}^{k} \phi_i(x) \geq 1$ for all x will be considered. The first requirement a selection rule must satisfy is that it has a certain minimum probability of selecting the best cell. Let P^*, $1/k < P^* < 1$, be a preassigned fixed number. Only rules which satisfy the _P*-condition_, viz.,

(2.1) $$\inf_{p} P_p(CS|\phi) \geq P^*,$$

will be considered. The set of rules which satisfy the P*-condition will be denoted by $\mathcal{D}_r^*$.

Two losses which are commonly used in subset selection problems are the size of the selected subset, S, and the number of non-best cells selected, S'. The risk using a rule ϕ when p is the true parameter is then the expected subset size, $E_p(S|\phi)$, or the expected number of non-best cells selected, $E_p(\tilde{S'}|\phi)$. Let

(2.2) $$S(x) = \sum_{i=1}^{k} \phi_i(x).$$

Then $S(x)$ is the expected subset size having observed $X = x$ and $E_p(S|\phi) = E_p(S(X)|\phi)$.

A selection rule, ϕ^*, is said to be <u>minimax with respect to</u> $\underline{S}$ if $\phi^* \in \mathcal{D}_{\Gamma^*}$ and

$$(2.3) \qquad \inf_{\mathcal{D}_p^*} \sup_{\underline{p}} E_{\underline{p}}(S|\underline{\phi}) = \sup_{\underline{p}} E_{\underline{p}}(S|\underline{\phi}^*).$$

Replace S by S' in (2.3) to define minimaxity with respect to S'. Berger (1979) investigated minimaxity with respect to S and S' in a general setting. In Section 3, selection rules which are minimax with respect to S and S' will be constructed.

3. CONSTRUCTION OF MINIMAX RULES

In Theorem 3.1, the following class of selection rules will be shown to be minimax with respect to S and S'. Let $\mathcal{D}^*$ be the class of selection rules which satisfy: (i) $\phi_i(\underline{x}) = \delta_i(x_i)$ for $1 \le i \le k$, i.e., the selection or rejection of the ith cell depends only on the number of observations in the ith cell; (ii) $\delta_i(m) \ge \delta_i(n)$ if $0 \le m \le n \le N$ and $1 \le i \le k$, i.e., the probability of selecting the ith cell decreases as the number of observations in the ith cell increases; (iii) $S(\underline{x}) = \sum_{i=1}^{k} \delta_i(x_i)$ is a Schur concave function of $\underline{x}$ (iv) $E_{\underline{p}_0}(\delta_i(X_i)) = P^*$ for $1 \le i \le k$ where $\underline{p}_0 = (1/k, \ldots, 1/k)$.

Remark 3.1. The requirement in (i) that the decision to include or exclude the ith cell depends only on X_i may seem to be a waste of the information contained in the other X_j's. But since $\sum_{i=1}^{k} X_i = N$, if X_i is "large" then the other X_j's must be

small and if X_i is "small", some of the other X_j's must be large. So information about the other X_j's is contained in X_i and the requirement is not as counterintuitive as it would be in a problem in which the X_j's are independent, e.g., the normal means problem (see Gupta (1965)). The rules considered by Nagel (1970) had property (i) as well as (ii) and (iv).

Remark 3.2. Condition (ii) seems reasonable since larger values of X_i indicate larger values of p_i.

Remark 3.3. See Proschan and Sethuraman (1977) and Nevius, Proschan and Sethuraman (1977) for a discussion of Schur functions in statistics. In a sense, $\underset{\sim}{x}$ majorizes $\underset{\sim}{y}$ if the coordinates of $\underset{\sim}{x}$ are more spread out than the coordinates of $\underset{\sim}{y}$. The more spread out the cell frequencies are, the easier it should be to decide which is the best cell and the smaller the subset that need be selected. This is an interpretation of condition (iii).

Remark 3.4. By Theorem 4.1 of Berger (1979), this is a necessary condition if a rule is to be minimax with respect to S or S'.

Remark 3.5. Gupta and Nagel (1967) proposed and studied the following selection rule for this problem: select the ith cell if $x_i \le x_{min} + C$ where $x_{min} = \min(x_1, \ldots, x_k)$ and C is a non-negative integer chosen so that the P*-condition is satisfied. In general, this rule is not in $\mathcal{D}^*$ since neither (i) nor (iii)

are true. In the special case of $k = 2$, however, it can be verified that the Gupta-Nagel rule is in $\mathcal{D}^*$ since $X_2 = N - X_1$. Of course, to satisfy (iv) exactly, this randomized version of the Gupta-Nagel rule must be used:

$$\phi_i(\underset{\sim}{x}) = \begin{cases} 1 & x_i < x_{min} + C \\ \alpha & x_i = x_{min} + C \\ 0 & x_i > x_{min} + C \end{cases}$$

So, by the following theorem, the Gupta-Nagel rule is minimax if $k = 2$.

Theorem 3.1. If $\underset{\sim}{\phi} \in \mathcal{D}^*$ then $\underset{\sim}{\phi}$ is minimax with respect to S and S'.

Proof. The first fact to be verified is that if $\underset{\sim}{\phi} \in \mathcal{D}^*$ then $\underset{\sim}{\phi}$ satisfies the P*-condition. Let $P_i = \{\underset{\sim}{p}: p_j \geq p_i \geq 0 \text{ for all } j \neq i \text{ and } \sum_{j=1}^{k} p_j = 1\}$. Then $\inf_{\underset{\sim}{p}} P_{\underset{\sim}{p}}(CS|\underset{\sim}{\phi}) = \inf_{1 \leq i \leq k} \inf_{P_i} P_{\underset{\sim}{p}}(\text{select } i^{th} \text{ cell}|\underset{\sim}{\phi}) = \inf_{1 \leq i \leq k} \inf_{P_i} E_{\underset{\sim}{p}}(\phi_i(\underset{\sim}{X}))$. Let Y have a binomial distribution with parameters N and p_i. Then $E_{\underset{\sim}{p}}(\phi_i(\underset{\sim}{X})) = E_{p_i}(\delta_i(Y))$. Since the binomial distribution has monotone likelihood ratio and since δ_i is a nonincreasing function, by Lemma 2, page 74 of Lehmann (1959), $E_{p_i}(\delta_i(Y))$ is nonincreasing in p_i. So the $\inf_{P_i} E_{\underset{\sim}{p}}(\phi_i(X)) = \inf_{P_i} E_{p_i}(\delta_i(Y)) = E_{p_i=1/k}(\delta_i(Y)) = P^*$. Thus $\inf_{\underset{\sim}{p}} P_{\underset{\sim}{p}}(CS|\underset{\sim}{\phi}) = \inf_{1 \leq i \leq k} P^* = P^*$ and $\underset{\sim}{\phi}$ satisfies the P*-condition.

Now (2.3) will be verified. By Theorem 3.1 of Berger (1979), the minimax value is kp^*. so it suffices to show that $\sup_{\underset{\sim}{p}} E_{\underset{\sim}{p}}(S|\underset{\sim}{\phi}) = kp^*$ for any $\underset{\sim}{\phi} \in \mathcal{D}^*$. But, since $S(\underset{\sim}{x})$ is Schur concave, by Application 4.2(a) of Nevius, Proschan and Sethuraman (1977) $E_{\underset{\sim}{p}}(S(\underset{\sim}{X})|\underset{\sim}{\phi})$ is Schur concave in $\underset{\sim}{p}$ and hence takes on the maximum value at $\underset{\sim}{p}_0 = (1/k, \ldots, 1/k)$.

$$E_{\underset{\sim}{p}_0}(S(\underset{\sim}{X})|\underset{\sim}{\phi}) = \sum_{i=1}^{k} E_{\underset{\sim}{p}_0} \phi_i(\underset{\sim}{X}) = \sum_{i=1}^{k} E_{\underset{\sim}{p}_0}(\delta_i(X_i)) = kP^* \text{ by (iv)}.$$

So (2.3) is verified.

Finally, by Theorem 3.2 of Berger (1979), if $\underset{\sim}{\phi}$ is minimax with respect to S then $\underset{\sim}{\phi}$ is also minimax with respect to S'. ||

Other authors have provided bounds on $E(S)$ for the rules they have proposed. For example, Gupta and Huang (1975) give an upper bound for $E(S)$ when the parameters are in a slippage configuration. But the result of Theorem 3.1 is stronger in that minimaxity considers all parameter configurations and the exact upper bound of kP^* for $E(S)$ is achieved.

$\mathcal{D}^*$ is a wide class of selection rules which are minimax. Finding one rule in $\mathcal{D}^*$ which has an additional optimality property may be accomplished by solving the following linear programming problem. Consider $\underset{\sim}{\delta} = (\delta_1(0), \ldots, \delta_1(N), \delta_2(0), \ldots, \delta_2(N), \ldots, \delta_k(0), \ldots, \delta_k(N))$ as the solution vector for which we wish to solve. Condition (ii) provides kN linear constraints on the solution. Condition (iii) provides additional linear constraints on the solution. For example, since (6, 2, 0) majorizes (5, 3, 1), we must have

$\delta_1(6) + \delta_2(2) + \delta_3(0) \leq \delta_1(5) + \delta_2(3) + \delta_3(1)$. Finally, condition (iv) provides k additional linear constraints on the solution. Subject to these constraints we wish to minimize $E_{\underset{\sim}{p}'}(S|\underset{\sim}{\delta})$, a linear function of the coordinates of $\underset{\sim}{\delta}$, for some $\underset{\sim}{p}' \neq \underset{\sim}{p}_0$. The parameter $\underset{\sim}{p}'$ could be some particular parameter value in which the experimenter is particularly interested. As an example, this problem was solved for $P^* = .4(.1).9$ and $\underset{\sim}{p}' = (.1, .45, .45)$ and $\underset{\sim}{p}' = (.2, .4, .4)$ in the case where $k = 3$ and $N = 6$. The resulting minimax rules are shown in Table 1. In finding these solutions, the additional constraint was made that the solution was permutation invariant, i.e., $\delta_i(m) = \delta_j(m)$ for all $1 \leq i, j \leq k$ and all $0 \leq m \leq N$. These solutions were obtained using the NYBLPC computer program. This program uses the criss-cross method as developed by Dr. Stanley Zionts. The computations were done on the CDC CYBER 74 at the Florida State University Computing Center.

4. A SIMPLE RULE

An examination of Table 1 reveals that, in all cases when P^* is large, the optimal rule obtained by solving the linear programming problem, as outlined in Section 3, has the following simple form;

$$(4.1) \qquad \phi_i^*(\underset{\sim}{x}) = \delta^*(x_i) = \begin{cases} 1 & x_i < t \\ \alpha & x_i = t. \\ 0 & x_i > t \end{cases}$$

Here t and α are chosen so that condition (iv) (Section 3) is satisfied. Namely, let Y have a binomial distribution with parameters 1/k and N. Then t is the integer which satisfies $P(Y < t) \leq P^* < P(Y \leq t)$ and $\alpha = (P^* - P(Y < t))/P(Y = t)$. This rule is analogous to a rule proposed by Nagel (1970) for selecting the largest cell probability. This rule is appealing for its simplicity and because the constants t and α, needed for implementation, can be easily obtained from a binomial table. In this section the rule ϕ^* is studied more closely. First, in Theorem 4.1, ϕ^* is shown to be minimax if P* is sufficiently large. Then, in Lemmas 4.1 and 4.2 and Theorem 4.2, ϕ^* is shown to have the following optimality property if the parameters are in a slippage configuration. Suppose all the cell probabilities are equal except for the ith. ϕ^* minimizes the probability of selecting the ith cell if its cell probability is larger than all the rest and maximizes the probability of selecting the ith cell if its cell probability is smaller than all the rest among all minimax rules.

Theorem 4.1. Let Y have a binomial distribution with parameters 1/k and N. Then, if

(4.2) $P^* \geq P(Y < \frac{N}{2}) + \frac{1}{2} P(Y = \frac{N}{2})$,

ϕ^* is minimax with respect to S and S'.

Proof. By Theorem 3.1 it suffices to show $\phi^* \in \mathcal{D}^*$. Conditions (i), (ii) and (iv) are obviously satisfied by the definition of

ϕ^*. It remains to show that (iii) is true, i.e., $S^*(\underset{\sim}{x}) =$ $\sum\limits_{i=1}^{k} \delta^*(x_i)$ is a Schur concave function of $\underset{\sim}{x}$.

To see that S^* is Schur concave, suppose $\underset{\sim}{x}$ majorizes $\underset{\sim}{y}$ where $\sum\limits_{\ell=1}^{k} y_\ell = N = \sum\limits_{\ell=1}^{k} y_\ell$. By the result of Hardy, Littlewood and Polya (1952, page 47), we may assume, without loss of generality, that $\underset{\sim}{x}$ and $\underset{\sim}{y}$ differ in two coordinates only, say, $x_i > y_i \geq y_j > x_j$ where $x_i + x_j = y_i + y_j$. Since the constants t and α are chosen so that $P(Y < t) + \alpha P(Y = t) = P^*$, (4.2) implies that either $t > \frac{N}{2}$ or $t = \frac{N}{2}$ and $\alpha \geq \frac{1}{2}$. Since all of the coordinates of $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are equal, except the ith and j^{th}, $S^*(\underset{\sim}{y}) - S^*(\underset{\sim}{x}) = \delta^*(y_i) + \delta^*(y_j) - \delta^*(x_i) - \delta^*(x_j)$. Since $\sum\limits_{\ell=1}^{k} y_\ell = N$ and $y_\ell \geq 0$, $1 \leq \ell \leq k$, $y_j \leq \frac{N}{2} \leq t$. If $y_j < t$ then $\delta^*(y_j) = \delta^*(x_j) = 1$ and $\delta^*(y_i) \geq \delta^*(x_i)$ so $S^*(\underset{\sim}{y}) - S^*(\underset{\sim}{x}) \geq 0$. If $t = y_j$ then $t = \frac{N}{2} = y_j = y_i$ and $S^*(\underset{\sim}{y}) - S^*(\underset{\sim}{x}) = \alpha + \alpha - 0 - 1 \geq 0$ since $\alpha \geq \frac{1}{2}$ if $t = \frac{N}{2}$. Thus S^* is Schur concave. ||

Values of the lower bound given in (4.2) are tabled in Table 2 for $k = 2(1) 10$ and $N = 1(1) 20$. Table 2 reveals that, for small values of k and N, this lower bound is reasonably small. But as k or N increases, the lower bound converges to one. So Theorem 4.1 shows that ϕ^* is minimax only for moderate values of k and N. It can be shown that if N is an odd number then the expression given in (4.2) is the same for both N and $N + 1$. That is why one column suffices for each consecutive odd-even pair in Table 2.

Now the behavior of ϕ^* when the parameter is in a slippage configuration will be examined. For the remainder of

this section let $p' = (p, \ldots, p, q, p, \ldots, p)$ where $(k - 1)p + q = 1$ and q is the ith coordinate. The following result will be proven.

Theorem 4.2. Suppose $\underset{\sim}{\phi}^*$ is minimax. Then among all minimax rules, $\underset{\sim}{\phi}$, $\underset{\sim}{\phi}^*$ maximizes $P_{\underset{\sim}{p}}$ (select i^{th} cell $|\underset{\sim}{\phi})$ if $q < p$ and $\underset{\sim}{\phi}^*$ minimizes $P_{\underset{\sim}{p}}$ (select i^{th} cell $|\underset{\sim}{\phi})$ if $q > p$.

The proof of Theorem 4.2 will be accomplished via the following two lemmas.

Lemma 4.1. Suppose $q < p$. Among all rules, $\underset{\sim}{\phi}$, which satisfy $P_{\underset{\sim}{p}_0}$ (select i^{th} cell $|\underset{\sim}{\phi}) \leq P^*$, $\underset{\sim}{\phi}^*$ maximizes $P_{\underset{\sim}{p}'}$ (select i^{th} cell $|\underset{\sim}{\phi})$.

Proof. For any $\underset{\sim}{p}$ and $\underset{\sim}{\phi}$, $P_{\underset{\sim}{p}}$ (select i^{th} cell $|\underset{\sim}{\phi}) = E_{\underset{\sim}{p}} \phi_i(\underset{\sim}{X})$. By the Neyman-Pearson Lemma, among all rules which satisfy $E_{\underset{\sim}{p}_0} \phi_i(\underset{\sim}{X}) \leq P^*$, a rule which maximizes $E_{\underset{\sim}{p}'} \phi_i(\underset{\sim}{X})$ is given by

$$(4.3) \qquad \phi_i(\underset{\sim}{x}) = \begin{cases} 1 & \\ \alpha & \frac{\binom{N}{\underset{\sim}{x}} q^{x_i} p^{N-x_i}}{\binom{N}{\underset{\sim}{x}} k^{-N}} \begin{array}{l} > t \\ \\ = \\ \\ < \end{array} \\ 0 & \end{cases}$$

where t and α are chosen so that $E_{\underset{\sim}{p}_0} \phi_i(\underset{\sim}{X}) = P^*$. Using the fact that $q < p$, it can be seen that (4.3) is equivalent to (4.1). $||$

Lemma 4.2. Let $q > p$. Among all rules, $\underset{\sim}{\phi}$, which satisfy $P_{\underset{\sim}{p}_0}$ (select i^{th} cell $|\underset{\sim}{\phi}) \geq P^*$, $\underset{\sim}{\phi}^*$ minimizes $P_{\underset{\sim}{p}}$ (select i^{th} cell $|\underset{\sim}{\phi})$.

Proof. The proof follows the same lines as that of Lemma 4.1. ||

Proof of Theorem 4.2. By Theorem 4.1 of Berger (1979), every rule, $\underset{\sim}{\phi}$, which is minimax with respect to S or S' must satisfy $P_{\underset{\sim}{p}_0}$ (select i^{th} cell $|\underset{\sim}{\phi}) = P^*$. So Theorem 4.2 follows from Lemmas 4.1 and 4.2. ||

Seal (1955, Section 4) examines a slippage problem, analogous to the one considered in Theorem 4.2, for the normal means selection problem. He shows, "approximately," that a certain rule has the property that $\underset{\sim}{\phi}^*$ has in the multinomial problem. "Approximately" is Seal's term. It refers to the fact that he used an asymptotic argument. But his result can be proved exactly using the Neyman-Pearson Lemma as in Lemmas 4.1 and 4.2.

5. SELECTION IN TERMS OF THE LARGEST CELL PROBABILITY.

In some problems the experimenter might be interested in selecting a subset of the cells including the cell associated with $p_{[k]}$, the largest cell probability. In this problem a correct selection, CS, is the selection of any subset including the cell associated with $p_{[k]}$. In this section, results analogous to those found in Sections 3 and 4 will be briefly outlined.

The following class of selection rules can be shown to be minimax with respect to S and S'. Let $\mathcal{D}_*$ be the class of selection rules which satisfy (i), (ii'), (iii) and (iv) where (i), (iii) and (iv) are as in Section 3 and (ii') $\delta_i(m) \leq \delta_i(n)$ if $0 \leq m \leq n \leq N$ and $1 \leq i \leq k$, i.e., the probability of selecting the ith cell increases as the number of observations in the ith cell increases.

Theorem 5.1. If $\phi \in \mathcal{D}_*$ then ϕ is minimax with respect to S and S'.

Proof. This proof follows the lines of the proof of Theorem 3.1 where now (ii') is used to show that if $\phi \in \mathcal{D}_*$ then ϕ satisfies the P*-condition. ||

Rules in $\mathcal{D}_*$ can be constructed by solving a linear programming problem. The condition that $\phi \in \mathcal{D}_*$ puts linear constraints on the vector of selection probabilities $\underset{\sim}{\delta} = (\delta_1(0), \ldots, \delta_k(N))$. The $\underset{\sim}{\delta}$ which minimizes $E_{\underset{\sim}{p}'}(S|\underset{\sim}{\delta})$ for some $\underset{\sim}{p}' \neq \underset{\sim}{p}_0$, subject to the constraint $\underset{\sim}{\delta} \in \mathcal{D}_*$ can be solved for. This was done for P* = .4(.1).9 and $\underset{\sim}{p}' = (.8, .1, .1)$ and $\underset{\sim}{p}'' = (.6, .2, .2)$ in the case where k = 3 and N = 6. The same rule was found to be optimal for both $\underset{\sim}{p}'$ and $\underset{\sim}{p}''$. The resulting minimax rule is presented in Table 3.

A particularly simple rule, which arose as the solution to the linear programming problem when P* was large, is the following:

$$(5.1) \qquad \phi_{*i}(\underset{\sim}{x}) = \delta_*(x_i) = \begin{cases} 1 & \text{if} \quad x_i > t \\ \alpha & \text{if} \quad x_i = t \\ 0 & \text{if} \quad x_i < t \end{cases}$$

where α and t are chosen so that condition (iv) is satisfied.
$\underset{\sim}{\phi}_*$ is the rule proposed by Nagel (1970). $\underset{\sim}{\phi}_*$, like $\underset{\sim}{\phi}^*$ in
Section 4, is minimax if P* is large and performs well in
the slippage configuration. Specifically, the following
two theorems are true.

<u>Theorem 5.1</u>. Let Y have a binomial distribution with parameters
$1/k$ and N. Then, if

$$(5.2) \qquad P^* \geq P(Y > 1) + \frac{1}{2} P(Y = 1),$$

$\underset{\sim}{\phi}^*$ is minimax with respect to S and S'.

<u>Proof</u>. If (5.2) is true then $\underset{\sim}{\phi}_* \in \mathcal{D}_*$. The key point to
verify is that $S_*(\underset{\sim}{x}) = \sum_{i=1}^{k} \delta_*(x_i)$ is Schur concave. This can be
verified using the fact that (5.2) implies either $t = 1$ and
$\alpha \geq \frac{1}{2}$ or $t = 0$. ||

For fixed values of N and P*, the lower bound given in
(5.2) is a decreasing function of k. So for fixed N and P*,
$\underset{\sim}{\phi}_*$ will be minimax if k is sufficiently large. The value of
the lower bound given in (5.2) is tabulated for N = 1 (1) 20
and k = 2 (1) 10 in Table 4.

Finally, $\underset{\sim}{\phi}_*$ performs well if the parameters are in a
slippage configuration. Let $\underset{\sim}{p}' = (p, \ldots, p, q, p, \ldots, p)$
where $(k - 1)p + q = 1$ and q is the ith coordinate. The
following result, analogous to Theorem 4.2, is true.

Theorem 5.2. Suppose ϕ_* is minimax. Then among all minimax rules, ϕ, ϕ_* maximizes P_Γ (select i^{th} cell $|\phi$) if $q > p$ and ϕ_* minimizes P_Γ (select $\tilde{i}^{th}$ cell $|\phi$) if $q < p$.

The proof of Theorem 5.2 is completely analogous to the proof of Theorem 4.2 and is omitted.

6. SELECTION RULES FOR POISSON PARAMETERS.

The selection rules constructed in Sections 3 and 5 can also be used for selection problems involving Poisson parameters. These rules will also be minimax for these Poisson selection problems. Selection rules for the Poisson distribution have previously been studied by Gupta and Nagel (1971) and Gupta and Huang (1975). Goel (1972) showed that the usual location and scale type selection rules do not satisfy the P*-condition in this problem for large values of P*.

Specifically let $\underset{\sim}{Y} = (Y_1, \ldots, Y_k)$ be independent Poisson random variables with parameters $\underset{\sim}{\lambda} = (\lambda_1, \ldots, \lambda_k)$. Let $N = \sum_{i=1}^{k} Y_i$. The goal is to select a subset of the population including the one associated with the largest or smallest parameter λ_i. The conditional distribution of $\underset{\sim}{Y}$ given $N = N$ is multinomial with parameters $\underset{\sim}{p}$ and N where $p_i = \lambda_i / \sum_{j=1}^{k} \lambda_j$. So if $N = N$, we can use the rule constructed in Section 3 or 5 for a sample of size N. This rule has the property that $P_{\underset{\sim}{\lambda}}(CS|N = N) \geq P^*$ for $0 \leq N < \infty$ and all $\underset{\sim}{\lambda}$. So unconditionally, $P_{\underset{\sim}{\lambda}}(CS) \geq P^*$ for all $\underset{\sim}{\lambda}$, using this rule, i.e., the rule satisfies the P*-condition. Similarly $E_{\underset{\sim}{\lambda}}(S) \leq kP^*$

for all $\underset{\sim}{\lambda}$ since $E_{\underset{\sim}{\lambda}}(S|N = N) \leq kP^*$ for $0 \leq N < \infty$ and all $\underset{\sim}{\lambda}$, i.e., the rule is minimax with respect to S. N is a complete sufficient statistic for $\underset{\sim}{\lambda}$ on the set $\Omega_0 = \{\underset{\sim}{\lambda}: \lambda_1 = \ldots = \lambda_k\}$. By Lemma 1.1 of Gupta and Nagel (1971) if $P_{\underset{\sim}{\lambda}}(CS|\underset{\sim}{\phi}) \equiv P^*$ on Ω_0, then $P_{\underset{\sim}{\lambda}}(CS|N = N) = P^*$ for $0 \leq N < \infty$. But $P_{\underset{\sim}{\lambda}}(CS|\underset{\sim}{\phi}) \equiv P^*$ on Ω_0 is a necessary condition for minimaxity by Theorem 4.1 of Berger (1979). So considering rules conditional on N is natural in this problem.

17

REFERENCES

Berger, R. L. (1979). Minimax subset selection for loss measured by subset size. (To appear Ann. Statist., 7, No. 6).

Goel, P. K. (1972). A note on the non-existence of subset selection procedures for poisson populations. Mimeograph Series No. 303, Department of Statistics, Purdue University.

Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. Technometrics, 7, 225-245.

Gupta, S. S. and Huang, D. Y. (1975). On subset selection procedures for Poisson populations and some applications to the multinomial selection problems. Applied Statistics (Gupta, R. P., ed.). North Holland Publishing Company.

Gupta, S. S. and Nagel, K. (1967). On selection and ranking procedures and order statistics from the multinomial distribution. Sankhya' Ser. B, 29, 1-34.

Gupta, S. S. and Nagel, K. (1971). On some contributions to multiple decision theory. Statistical Decision Theory and Related Topics. (Gupta, S. S. and Yackel, J. eds.). Academic Press.

Hardy, G. H., Littlewood, J. E. and Polya, G. (1952). Inequalities. Cambridge University Press.

Lehmann, E. L. (1959). Testing Statistical Hypotheses. John Wiley and Sons.

Nagel, K. (1970). On subset selection rules with certain optimality properties. Ph.D. Thesis, Purdue University.

Nevius, S. E., Proschan, F. and Sethuraman, J. (1977). Schur functions in statistics II. Stochastic majorization. Ann. Statist., 5, 263-274.

NYBLPC. State University of New York at Buffalo Computing Center Press.

Panchapakesan, S. (1971). On a subset selection procedure for the most probable event in a multinomial distribution. Statistical Decision Theory and Related Topics. (Gupta, S. S. and Yackel, J., eds.). Academic Press.

Proschan, F. and Sethuraman, J. (1977). Schur functions in statistics I. The preservation theorem. Ann. Statist., 5, 256-263.

Seal, K. C. (1955). On a class of decision procedures for ranking the means of normal populations. Ann. Math. Statist., 26, 387-398.

Table 1

Minimax selection rule for $p_{[1]}$ which minimizes $E_{\varrho'}(S)$

$$\varrho' = (.1, .45, .45) \quad N = 6 \quad k = 3$$

$$\delta(x) = P(\text{select } i^{th} \text{ cell} \mid X_i = x)$$

x \ P*	.9	.8	.7	.6	.5	.4
0	1.000	1.000	.886	.759	.633	.506
1	1.000	1.000	.886	.759	.633	.506
2	1.000	1.000	.886	.759	.633	.506
3	1.000	.545	.443	.380	.316	.253
4	.002	.000	.000	.000	.000	.000
5	.000	.000	.000	.000	.000	.000
6	.000	.000	.000	.000	.000	.000
$E_{\varrho'}(S\mid\delta)$	2.489	2.206	1.929	1.654	1.378	1.102

$$\varrho' = (.2, .4, .4)$$

x \ P*	.9	.8	.7	.6	.5	.4
0	1.000	1.000	.778	.667	.556	.500
1	1.000	1.000	.778	.667	.556	.500
2	1.000	1.000	.778	.667	.556	.500
3	1.000	.545	.778	.667	.556	.272
4	.002	.000	.000	.000	.000	.000
5	.000	.000	.000	.000	.000	.000
6	.000	.000	.000	.000	.000	.000
$E_{\varrho'}(S\mid\delta)$	2.625	2.336	2.042	1.750	1.458	1.168

Table 2

Lower bound for P*

If P* ≥ table entry, ȼ* is minimax.

N k	1,2	3,4	5,6	7,8	9,10	11,12	13,14	15,16	17,18	19,20
2	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000
3	.6667	.7407	.7901	.8267	.8552	.8779	.8965	.9118	.9245	.9352
4	.7500	.8438	.8965	.9294	.9511	.9657	.9757	.9827	.9876	.9911
5	.8000	.8960	.9421	.9667	.9804	.9883	.9930	.9958	.9974	.9984
6	.8333	.9259	.9645	.9324	.9910	.9954	.9976	.9987	.9993	.9996
7	.8571	.9446	.9767	.9398	.9955	.9979	.9991	.9996	.9998	.9999
8	.8750	.9570	.9839	.9938	.9975	.9990	.9996	.9998	.9999	1.0000
9	.8889	.9657	.9885	.9960	.9986	.9995	.9998	.9999	1.0000	1.0000
10	.9000	.9720	.9914	.9973	.9991	.9997	.9999	1.0000	1.0000	1.0000

Table 3

Minimax selection rule for $p_{[k]}$ which

minimizes $E_{\underline{p}'}(S)$ and $E_{\underline{p}''}(S)$

$\underline{p}' = (.8, .1, .1)$ $\underline{p}'' = (.6, .2, .2)$ $N = 6$ $k = 3$

$\delta(x) = P(\text{select } i^{\text{th}} \text{ cell } |X_i = x)$

x \ p*	.9	.8	.7	.6	.5	.4
0	.000	.000	.000	.000	.000	.000
1	.954	.574	.448	.384	.320	.253
2	1.000	1.000	.897	.768	.639	.507
3	1.000	1.000	.897	.768	.639	.507
4	1.000	1.000	.897	.768	.639	.507
5	1.000	1.000	.897	.768	.680	.747
6	1.000	1.000	1.000	1.000	1.000	1.000
$E_{\underline{p}'}(S|\delta)$	1.904	1.635	1.446	1.276	1.122	1.025
$F_{\underline{p}''}(S|\delta)$	2.433	2.121	1.852	1.594	1.341	1.112

Table 4

Lower bound for P*

If P* ≥ table entry, $\mathfrak{L}_*$ is minimax

k	2	3	4	5	6	7	8	9	10
N									
1	.2500	.1667	.1250	.1000	.0833	.0714	.0625	.0556	.0500
2	.5000	.3333	.2500	.2000	.1667	.1429	.1250	.1111	.1000
3	.6875	.4815	.3672	.2960	.2477	.2128	.1865	.1660	.1495
4	.8125	.6049	.4727	.3856	.3248	.2803	.2463	.2196	.1981
5	.8906	.7037	.5649	.4675	.3972	.3446	.3039	.2717	.2455
6	.9375	.7805	.6440	.5412	.4642	.4051	.3589	.3218	.2914
7	.9648	.8390	.7108	.6068	.5256	.4618	.4110	.3697	.3357
8	.9805	.8829	.7664	.6645	.5814	.5144	.4600	.4154	.3782
9	.9893	.9155	.8123	.7148	.6318	.5630	.5061	.4587	.4189
10	.9941	.9393	.8498	.7584	.6770	.6076	.5490	.4996	.4576
11	.9968	.9566	.8803	.7960	.7174	.6483	.5889	.5381	.4944
12	.9983	.9692	.9050	.8262	.7533	.6855	.6259	.5742	.5293
13	.9991	.9782	.9248	.8557	.7850	.7192	.6601	.6080	.5622
14	.9995	.9846	.9406	.8791	.8131	.7497	.6916	.6395	.5933
15	.9997	.9892	.9532	.8988	.8377	.7772	.7205	.6689	.6225
16	.9999	.9924	.9633	.9156	.8594	.8019	.7470	.6962	.6500
17	.9999	.9947	.9712	.9296	.8783	.8242	.7712	.7215	.6757
18	1.0000	.9963	.9774	.9415	.8948	.8441	.7934	.7450	.6998
19	1.0000	.9974	.9824	.9514	.9092	.8619	.8136	.7666	.7223
20	1.0000	.9982	.9863	.9596	.9217	.8778	.8319	.7866	.7433

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
FSU No. M497 USARO-D 55		

4. TITLE	5. TYPE OF REPORT & PERIOD COVERED
Minimax Subset Selection for the Multinomial and Poisson Distributions	Technical Report
	6. PERFORMING ORG. REPORT NUMBER
	FSU Statistics Report M497

7. AUTHOR(S)	8. CONTRACT OR GRANT NUMBER(S)
Roger L. Berger	USARO Grant No. DAAG29-76-G-0238

9. PERFORMING ORGANIZATION NAME & ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
The Florida State University Department of Statistics Tallahassee, Florida 32306	

11. CONTROLLING OFFICE NAME & ADDRESS	12. REPORT DATE
U.S. Army Research Office-Durham P.O. Box 12211 Research Triangle Park, N.C. 27709	February, 1979
	13. NUMBER OF PAGES
	21

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS (of this report)
	Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (Of this report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (Of the abstract entered in Block 20, if different from report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS

Linear programming, minimax subset selection, expected subset size.

20. ABSTRACT

Let $(X_1, \ldots, X_k)$ be a multinomial vector with unknown cell probabilities $(p_1, \ldots, p_k)$. A subset of the cells is to be selected in a way so that the cell associated with the smallest cell probability is included in the selected subset with a preassigned probability, P^*. Suppose the loss is measured by the size of the selected subset, S. Using linear programming techniques, selection rules can be constructed which are minimax with respect to S in the class of rules which satisfy the P^*-condition. In some situations, the rule constructed by this method is the rule proposed by Nagel (1970) Similar techniques also work for selection in terms of the largest cell probability. The rules constructed in this fashion are also minimax for selection in terms of Poisson parameters.